# METHOD AND APPARATUS FOR INDUCING CLASSIFIERS FOR MULTIMEDIA BASED ON UNIFIED REPRESENTATION OF FEATURES REFLECTING DISPARATE MODALITIES

## ABSTRACT

5  This invention is a system and method to perform categorization (classification) of multimedia items. These items are comprised of a multitude of disparate information sources, in particular, visual information and textual information. Classifiers are induced based on combining textual and visual feature vectors. Textual features are the traditional ones, such as, word count vectors. Visual features include, but are not limited to, color properties of key intervals and motion

10  properties of key intervals. The visual feature vectors are determined in such a fashion that the vectors are sparse. The vector components are features such as the absence or presence of the color green in spatial regions and the absence or the amount of visual flow in spatial regions of the media items. The text and the visual representation vectors are combined in a systematic and coherent fashion. This vector representation of a media item lends itself to well-established

15  learning techniques. The resulting system, subject of this invention, categorizes (or classifies) media items based both on textual features and visual features.